Strategic Evolution of ESE Data Systems (SEEDS)
Formulation Study
Overview

June 17, 2002

Presented By

Dr. *Stephen Wharton*   *SEEDS Formulation Manager*
301-614-5350* *wharton@eosdata.gsfc.nasa.gov

SEEDS Formulation Team

Catherine Corlan, Kathy Fontaine, Vanessa Griffin,
Gail McConaughy, Ken McDonald, Karen Moe, H. Ramapriyan,
Matt Schwaller, Richard Ullman, Stephen Wharton

❑ Charter:

- ➢ Using the NewDISS Strategy Document as pre-formulation concept, establish a strategy for the evolution of the Earth Science Enterprise network of data systems and service providers over the next decade that:
  - ▪ Ensures the timely delivery of Earth Science information at an affordable cost.
  - ▪ Fully engages the community on data management issues, objectives, and solutions.

❑ Scope (as traced to three questions from the NewDISS Strategy Document):

- ➢ How to introduce greater flexibility and responsiveness into the standards, processes, and infrastructure used to support the generation of science data products from NASA science missions?
  - ▪ This **Unifying Framework** is the principal focus of the formulation effort.
- ➢ How to establish an integrating framework of data management standards, processes, and infrastructure that leverages information technologies that exist or are being developed by other agencies/commercial sector?
  - ▪ Plans and processes for **Technology Infusion** to be coordinated with ESTO.
- ➢ How to establish a more broadly based network of science and applications providers of products, system capabilities, and services to address the ESE strategic science objectives?
  - ▪ SEEDS expects to play a supporting role in **Program/Project Management** by assisting NASA HQ in developing specifications for AO/NRA's and in evaluating the proposals received.
  - ▪ SEEDS will survey program/project governance approaches that fully engaged the community in establishing a broadly based network of science and applications providers.
  - ▪ SEEDS does not expect to be responsible for contract support, implementation, operation, or data center management.

❏ **Initiate Formulation Studies.**

➢ Establish contract support tasks to provide a coordination and supporting role.

➢ Leverage community expertise as consultants and/or volunteers.

❏ **Assess where we are today:**

➢ Use NewDISS Strategy Document as pre-formulation concept.

➢ Study teams survey existing practices, lessons learned.

➢ Identify science concerns / issues pertinent to the study.

➢ Hold public workshops.

❏ **Formulate recommendations:**

➢ Address findings from assessment.

➢ Revise/refine survey, questions, concerns, recommendations in response to community feedback.

➢ Make surveys, preliminary findings, reports, and recommendations available via web.

➢ Integrate study results into policy guidelines.

➢ Recommend roles and responsibilities for SEEDS Office.

➢ Present recommendations and plans to ESISS, ESSAAC, and NRC.

➢ Submit recommendations by July 2003 to Enterprise AA for approval.

❏ **Initiate planning and analysis efforts to prepare for execution phase in FY04.**

❑ **The primary formulation results will be in the form of recommendations to the Earth Science Enterprise (sub-bullets noted below).**

❑ **To address the recommendations from the NewDISS Strategy Document the Formulation Team has established seven study teams.**

➢ Cost Estimation and Levels of Service – Vanessa Griffin
  ▪ Recommend level of service guidelines.
  ▪ Develop cost model and benchmarks.

➢ Standards and Interfaces for Near Term Missions – Richard Ullman
  ▪ Recommend standards and interfaces for missions starting development.

➢ Standards and Interfaces for Future Missions – Ken McDonald
  ▪ Recommend process to establish standards and interfaces for future (2004+) missions.

➢ Life Cycle Data Management for Long Term Archive – Matt Schwaller
  ▪ Recommend guidelines for data lifecycle and LTA planning.

➢ Software Reuse and Reference Architecture – Gail McConaughy
  ▪ Develop options to promote reuse reduce cost and risk.

➢ Metrics Planning and Reporting – H. Ramapriyan
  ▪ Recommend metrics to monitor results from data system and service providers.
  ▪ Recommend options for governance that more fully engage the community.

➢ Technology Needs and Infusion Planning – Karen Moe:
  ▪ Identify future data management technology infusion needs/opportunities.

- **The community is involved at many levels in SEEDS Formulation:**
  - As participants and consultants on the various study teams.
  - As contributors of white papers, workshop attendees, and survey respondents.
  - As advisory panels in reviewing SEEDS plans and recommendations (ESISS, ESSAAC, NRC).

- **Study Teams:**
  - Initial surveys completed.
  - Shared their status and discussed their efforts with the community.
  - Working on preliminary ideas for recommendations.

- **Implemented SEEDS Public Web Page at http://eos.nasa.gov/seeds.**

- **Held first Public Workshop (Feb 5-7, 2002).**
  - Significant participation from the data provider community.
  - Received feedback on additional "process" elements to be considered.
  - Received 15 white papers and 40 cost team recommendations.
  - "Signed up" three community volunteers for metrics reporting and life cycle study team.

- **Second Public Workshop (June 17-19, 2002):**
  - Workshop focus is "processes" for the SEEDS era.
  - Plan to solicit deeper involvement by community through:
    - presentations on "best practices" in other environments (NASA and non-NASA) and
    - breakout sessions focused on on-going "deep involvement" by community.

❑ ESE Enterprise:

- ➢ Sets requirements, priorities, and budget.
- ➢ Issues calls for proposals, evaluates proposals, and awards funding to selected teams.
- ➢ Selects appropriate NASA Program(s)/Project(s) to oversee development and operation.

❑ SEEDS "Office" (ideas for execution phase):

- ➢ Conduit for community involvement in formulating ESE data management policies and standards.
- ➢ With community, establishes and evolves framework of standards, levels of service, and interfaces.
- ➢ Utilize AO/BAA/CAN mechanism to fund broad community participation in:
  - ▪ Supporting definition and effective use of standards and interfaces for data and information systems.
  - ▪ Establishing reuse architecture working groups.
  - ▪ Supporting infusion of advanced technology and standards into currently existing ESE data systems.
- ➢ Provide guidelines for ESE data system related AO's and NRA's.
- ➢ Assists ESE in evaluating proposals and compiles performance metrics of data services providers.
- ➢ Assists programs, projects, or implementing teams in applying SEEDS recommendations.

❑ NASA Implementing Program/Project:

- ➢ Establish requirements, plans, schedules, deliverables with the selected implementing team(s).
- ➢ Manages funding, monitors progress, and ensures compliance with standards and requirements.
- ➢ Provides data management infrastructure support as appropriate.

❑ Implementing Team:

- ➢ Defines science and/or applications requirements for products and services.
- ➢ Develops approach, selects partners, and carries out design, implementation, and operation.

❑ **Challenges for SEEDS:**

  ➢ *How do you involve the broader community during these initial and following phases: including both users and providers?*

  ➢ *How does one decide what falls within the scope of SEEDS and what falls outside?*

  ➢ *How does SEEDS deal with rapidly changing user requirements?*

  ➢ *How do the present plans relate to the national and international dimensions?*

❑ **Recommendations:**

  ➢ The scope of SEEDS:

    ▪ At one extreme the scope of SEEDS is merely a set of interfaces and standards with some costing tools. As an activity this may fairly characterize it at present.

    ▪ But implicitly or explicitly in many discussions the various types of functions, processes and data centers are also included and hence this concept of SEEDS represents a much more comprehensive concept.

    ▪ The original concept of NewDISS identified interfaces and standards as the principal components that had to be well defined. This did not imply that this was all that NewDISS should be and the early conceptual documents on NewDISS included all the various data centers as well, though it was concluded that the number and the specific functions of these did not need initially to be closely detailed.

    ▪ But within NASA there will need to be an ability to monitor performance and to make strategic decisions concerning the allocation of resources between different types of processes, functions and data centers.

  ➢ Current work by the Formulation Team represents a major advance in the planning of SEEDS but to be successful it must engage user and provider communities much more closely.

  ➢ It should be the highest priority for the current Formulation Team of the SEEDS activity to develop and implement organizational structures facilitating much deeper engagement of key stakeholders. This action itself must involve some of these stakeholders and should start immediately.

❑ **The focus of this workshop will be on developing community-based processes to guide the development and validation of SEEDS protocols and standards, cost models, interfaces, technology, and accountability.**

❑ **Plenary Sessions:**
  ➢ Welcoming Remarks   Steve Wharton, GSFC
  ➢ Welcoming Remarks   Martha Maiden, NASA HQ
  ➢ Using Data Grid Technology to Manage Distributed Data   Reagan Moore, UCSD/SDSC
  ➢ Deep Community Involvement   John Townshend, UMD
  ➢ ESIP Federation Insights on Technology Findings   Rob Raskin, JPL
  ➢ Capability/Technology Study Team   Karen Moe, GSFC

❑ **Concurrent Breakout Sessions:**
  ➢ Capability/Technology 1 & 2
    ▪ The objectives of the SEEDS technology study are to define and conduct community-based processes to identify needed technical capabilities in the SEEDS era, and to define technology infusion processes.
    ▪ This session will engage the ESE community to help in the formulation and design of a SEEDS capabilities vision. Participation from the full spectrum of data system users, developers and managers is desired
  ➢ Community Involvement 1 & 2
    ▪ The goal of the breakout session on community involvement is to gather strategies for bringing various community elements together towards a common outcome.
    ▪ This session seeks to engage the participants in discussions based on the issues presented this morning.

❑ **Plenary Sessions:**

➢ Breakout Summary - Capability/Technology	Karen Moe, GSFC
➢ Breakout Summary - Community Involvement	Kathy Fontaine, GSFC
➢ ESISS/ESSAAC Report	Sara Graves, UAH
➢ Standards Issues	Sylvia Nittel, University of Maine
➢ Preserving Earth Science Data for the Future	Graham Bothwell, JPL
➢ Standards Processes and Near-Term Recommendations Report	Ken McDonald, GSFC
➢ Life Cycle Study Team Report	Matt Schwaller, GSFC

❑ **Concurrent Breakout Sessions:**

➢ Standards 1 & 2

  ▪ Objectives - To get user community's feedback on near-term recommendations for data, metadata, and data interface standards and for processes for evolution of standards for ESE.

➢ Lifecycle

  ▪ Objective - Review SEEDS Lifecycle Study Team results with an invited panel of experts representing spacecraft and instrument data products, science data processing, active and long-term archive.

Plenary Sessions:

- Breakout Summary – Standards   Ken McDonald /Rich Ullman, GSFC
- Breakout Summary - Life Cycle   Matt Schwaller, GSFC
- Reuse Lessons Learned   Ed Masuoka, GSFC
- Creating Community Involvement in Interoperability, Metrics, and Technology Development: The Case of the ESIP Federation   Bruce Caron, ESIP Federation
- Reuse   Gail McConaughy, GSFC
- Metrics   Rama Ramapriyan, GSFC

Concurrent Breakout Sessions:

- Reuse
  - Goal - Define a process to enable software reuse.
- LOS/Cost Model
  - The Level of Service/Cost Estimation team will discuss progress to date on cost estimation by analogy.
- Metrics 1 & 2
  - Goal - Obtain ideas from the community regarding processes for metrics planning and reporting, accountability and governance mechanisms that can be used in the Earth Science Enterprise's data system environment of the future.

Plenary Sessions:

- Closing Remarks   Vanessa Griffin, GSFC

Acronym List

AA	Associate Administrator
AO	Announcement of Opportunity
COTS	Commercial Off the Shelf
DAAC	Distributed Active Archive Center
ECS	EOSDIS Core System
EDC	Eros Data Center
EOSDIS	Earth Observing System Data Information System
ESISS	ESSAAC Subcommittee on Information Systems and Services
ESE	Earth Science Enterprise
ESIP	Earth Science Information Partnership
ESSAAC	Earth System Science and Applications Advisory Committee
ESTO	Earth Science Technology Office
FT	Formulation Team
FTE	Full Time Equivalent
FY	Fiscal Year
GPM	Global Precipitation Mission
GSFC	Goddard Space Flight Center
GST	Global Science & Technology
ISO	International Standards Organization
LaRC	Langley Research Center
LOS	Level of Service
LTA	Long Term Archive
NBC	New Business Committee
NewDISS	New Data and Information Systems and Services
NOAA	National Oceanic and Atmospheric Administration
NRC	National Research Council
OGC	Open GIS Consortium
ORNL	Oak Ridge National Laboratory
PR	Procurement Request
RFP	Request for Proposal
SEEDS	Strategic Evolution of ESE Data Systems
SGT	Stinger Ghaffarian Technologies
SOW	Statement of Work
SSAI	Science Systems and Applications Inc.
UWG	User Working Group